12-Dec-2018

## Syllabus

## Paper I

## unit 3:- Research Methods

1) Statistics → Relevance, Basic concepts, Terminology, Descriptive & Inferential Statistics

2) Hypothesis ~~Theor~~ Testing, Normal Distribution, Type I, Type II errors, z-Test, t-test, etc

3) Correlation & Regression

4) ANOVA

5) Factor Analysis

6) Research Designs
(ex post facto & Experimental)

7) Item Response Theory (IRT)

8) Signal Detection (SDT)

9) Sampling Type, Errors, Scale of Measurement

10) Threshold

⇒ Research is relevant because sample keeps changing with time.

⇒ Statistics are manipulative in nature → small changes are not noticed but they aggregate to a bigger change over time

Q) What is Research?

Ans:- Roughly speaking, research refers to search of knowledge. Technically, research is a scientific and systematic search for pertinent information on a specific topic. It is an art of scientific investigation which involves defining & re-defining problems, formulating hypothesis or suggested solutions, collecting, organising & evaluating the data

SHOUT
SHARE
IT

making deductions & conclusions and finally testing the conclusions to see whether they fit the formulated hypothesis.

There are diff. kinds of researches like descriptive, analytical, applied, fundamental, etc.

Q) Elaborate the steps in doing a research (in general)?

Q) Elaborate the steps performed during a research in behavioural sciences.

* fundamental Research → fundamental to humans
eg. genetics

Q) Find out the diff. types of research & elaborate them giving suitable examples (Differentiate between them)

The 2 parameters that any test or measure should fulfill are validity & reliability.

## Role of statistics

The role of statistics in research is to function as a tool in designing research, analysing its data and drawing conclusions. Most research studies result in a large volume of raw data which must be suitably reduced so that the same can be used for further analysis. It is an art of drawing conclusion about phenomena in which chance plays a role. eg. predicting the state of weather.
The randomness may occur due to:

(i) Intrinsic random nature of the phenomena

(ii) Unavoidable Noise in the Environment

(iii) Conscious randomisation of experimental or measurement units

## Relevance in Behavioural Sciences

(i) It leads to quantification of the data, that means assigning numerical values to abstract & subjective concepts.

(ii) Psychologists use statistics because it allows for the establishment of models and empowers them to predict behaviour based on the findings from the data. eg. Dyslexia treatment

(iii) From the development of new therapy techniques to evaluating the effectiveness of the techniques upon implementation, it is statistical analysis which provides the means by which conclusions can be drawn.

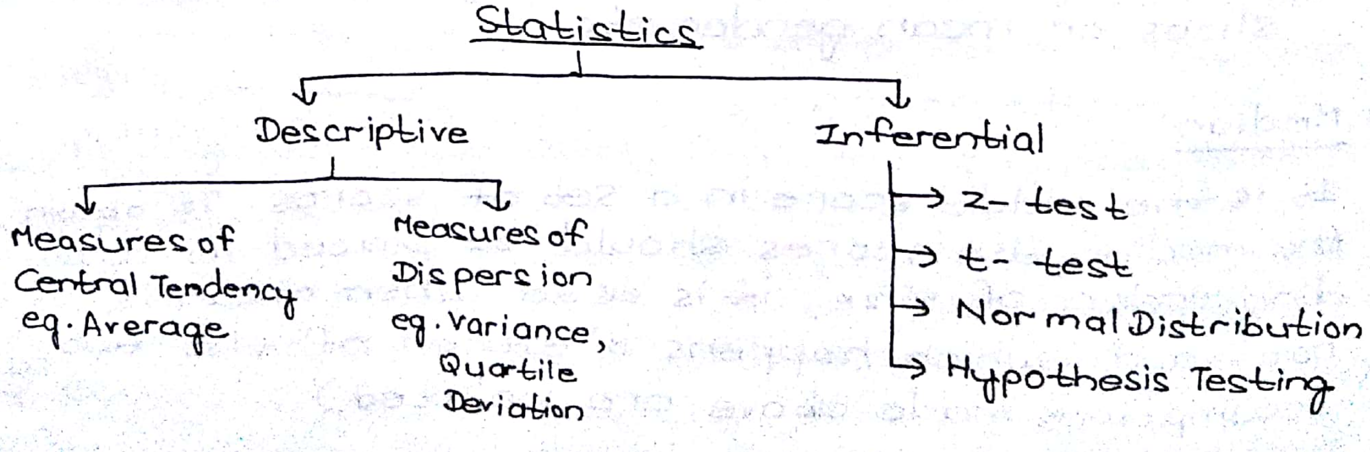
(iv) To answer questions like – How do changes in one variable impact other variables? Is there a way we can measure ~~a~~ the relationship ~~& rela~~ between diff. variables? What is the overall strength of the relationship & what does it mean?

## Terminology –

(i) Population :- It refers to all members of some pre-defined group also known as universe. It usually has a large no. of participants (taken to be infinite)

(ii) Sample :- It is a subset of the population on which study is performed. It is to be chosen with utmost care using sampling techniques. It should be random & unbiased so that it leads to the correct representation of the population.

There are 2 major areas of study in statistics.


Statistics
Descriptive
Inferential
Measures of Central Tendency eg. Average
Measures of Dispersion eg. Variance, Quartile Deviation
z- test
t- test
Normal Distribution
Hypothesis Testing

## *Measures of central tendency

(i) Mean / Average
(ii) Median
(iii) Mode

13-Dec-2018

## * Mean / Average / Arithmetic Mean

Let '$x$' be the 'score'
'$n$' = total no. of subjects

$$\boxed{\bar{x} = \frac{\Sigma x}{n}}$$

frequency (f)

$$\boxed{\bar{x} = \frac{\Sigma xf}{\Sigma f}}$$

eg.

| $x$ | $f$ | $fx$ |
|---|---|---|
| 10 | 20 | 200 |
| ⋮ | ⋮ | ⋮ |
|  | $\Sigma f$ | $\Sigma fx$ |

* Narrower the research better it is because the general populations vary a lot, so research should be specific.

Calculating avg. is specially useful in questions pertaining to temp., IQ, time taken to travel, etc. When we calculate & interpret the mean, we are required to make 2 assumptions about our data:

(i) The distribution is symmetrical
(ii) The data is measured at the interval or ratio scale eg. there is no meaning of avg. colour of shoes or mean gender, etc.

## * Median

It is the middle score in a set of scores. To obtain the median, the scores should be placed in ascending order of size, it is used when mean is not valid (which happens if either of the two assumptions made above are violated)

Median Location $= \frac{N+1}{2}$ (if 'N' is odd)

$= $ Avg. of score at $\frac{N}{2} + \left(\frac{N+1}{2}\right)^{th}$ position

(if 'N' is even)

*IQ scores of 5 people were 93, 81, 81, 95 and 200
Avg:- 110 & Median:- 93

* In most cases, mean is better as it takes contribution of all members into account.

- The median is sometimes used when a set of scores include one or two scores that are very different from the rest, when this outcome occurs, mean gives a distorted value.

* Mode

It is defined as the most frequent score in the distribution. The data can be unimodal or bimodal or multimodal.

* When the distribution is unimodal & symmetrical, then the mode, median & mean will have very similar values.

* The mean is often the best avg. because unlike the median, it uses all the information available i.e. every member in the dataset has an influence on the avg. It also has useful distributional properties. However, it is affected by skew. When the distribution is skewed, it has an effect of pulling the mean away from the true value.

*Measures of dispersion

central tendency doesn't mean a lot without a measure of dispersion or spread.

(1) * Range

The difference bet. highest & lowest scores of a group

## * Quartile

(i) First Quartile ($Q_1$) :- $\frac{1}{4}$th way up
(25th centile)

(ii) Second Quartile ($Q_2$) :- Half way up
(Median / 50th centile)

(iii) Third Quartile ($Q_3$) :- $\frac{3}{4}$th way up
(75th centile)

IQR :- Inter Quartile Range

$\boxed{Q_3 - Q_1}$ (Upper Quartile – Lower Quartile)

## (2) Variance & Standard Deviation

$$\boxed{\sigma^2 = \frac{\Sigma (x_i - \bar{x})^2}{N}}$$

$$\boxed{\sigma = \sqrt{\frac{\Sigma (x_i - \bar{x})^2}{N}}}$$

- Variance is the avg. of square of distance of the scores from the mean. It is a measure of the avg. amount by which the scores in the sample deviate from the mean score. It is the most popular measure of the scatterdness of the data.

- Standard Deviation ($\sigma$) :- square root of variance (only +ve root considered)

## * (3) Mean Deviation

$$\boxed{\text{M.D. about } \bar{x} = \frac{\Sigma |x_i - \bar{x}|}{N}}$$

$$\boxed{\text{M.D. about a point 'a'} = \frac{\sum |x_i - a|}{N}}$$

## Inferential Statistics

### (1) Normal Distribution

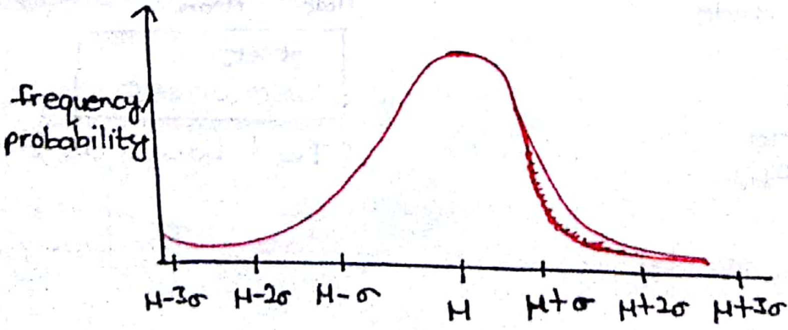
↳Carl Gauss

It is the background of inferential statistics developed by carl Gauss in the 18th century.

We say 'x' follows normal distribution with '$\mu$' & '$\sigma$'

$\boxed{X \sim N(\mu, \sigma)}$ ($\sigma$ is +ve)

A normal distribution is symmetrical about y-axis & bell shaped. It is thick in the middle and tails get thinner and thinner (asymptotic). A very large ~~of~~ no. of naturally occuring variables are normally distributed. The reason why everyone tries to approximate the distribution as normal is that we have mathematical equations, normal tables, etc. governing it. A lot of statistical tests make the assumption that data is normally distributed. It is a unimodal distribution. In fact, mean, median & mode co-incide, it comes close to fitting many observed data like weight, age, height, IQ, agricultural output, learning performance, etc.



The great advantage of a normal distribution is that if you know or can estimate 2 values or parameters i.e. mean & standard deviation, then we can calculate all possible probabilities as well as other stat. quantities. We can use area under the curve to find the probability of any range of values.

Meaning of one standard deviations of the mean is the values bet. '$\mu - SD$' ~~or~~ & '$\mu + SD$'

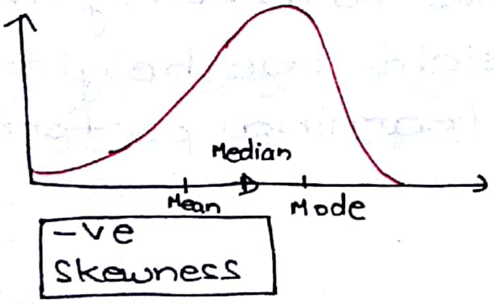
- 2 SDs of the mean :- values bet. $\mu - 2\sigma$ & $\mu + 2\sigma$

- 3 SDs of the mean :- values bet. $\mu - 3\sigma$ & $\mu + 3\sigma$

*We know that 68 % of the area is bet. 1SDs of the mean, 95 % of the area bet. 2SDs and 99.7 % area bet. 3SDs.

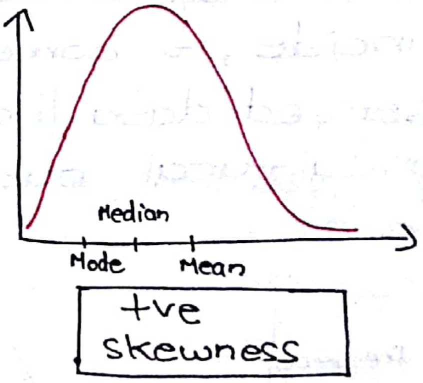
- Skewness

It is a measure of asymmetry and shows the manner in which the items are clustered around the avg. In symmetric distributions, skewness = 0 but if the curve is distorted, we have asymmetrical distribution & it indicates the presence of skewness. The distortion can be on the right or the left side.



(Tail towards the left)

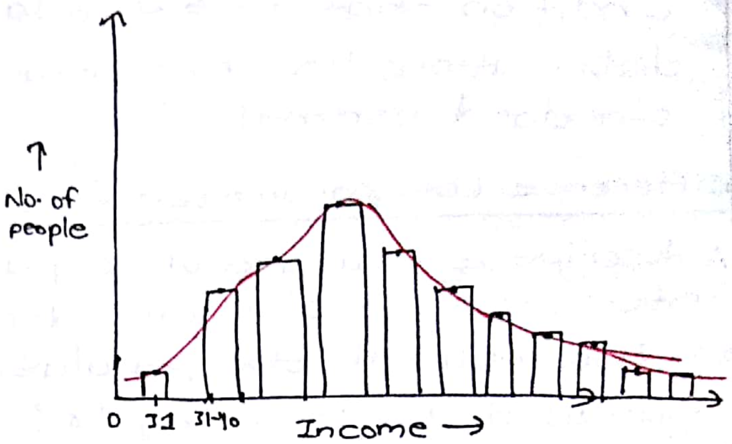
Mean < Median < Mode



(Tail towards the right)

Mode < Median < Mean

| * Frequency | | Income |
|---|---|---|
| 3 | → | <31 |
| 50 | → | 31-40 |
| 70 | → | 41-50 |
| 87 | → | 51-60 |
| 70 | → | 61-70 |
| 40 | → | 71-80 |
| 30 | → | 81-90 |
| 20 | → | 91-100 |
| 11 | → | 101-110 |
| 9 | → | 111-120 |
| 5 | → | 121-130 |
| 3 | → | 131-140 |
| 2 | → | 141-150 |



(+vely skewed data)

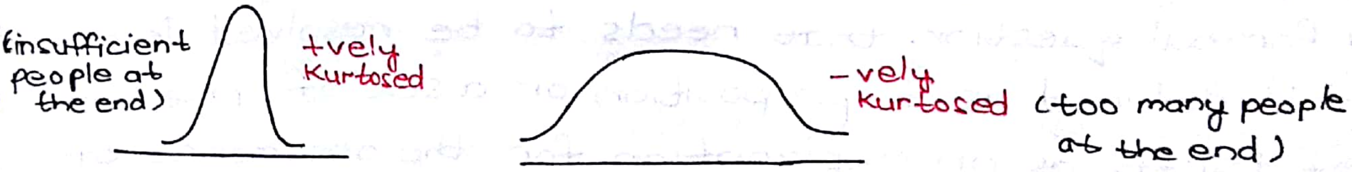
*Pearson's coefficient of skewness

$$\boxed{\frac{\text{Mean} - \text{Mode}}{\text{S.D.}}}$$

(+ve for positive skewness & -ve for negative skewness)

15-Dec-2018

* Kurtosis


(insufficient people at the end)
+vely Kurtosed
-vely Kurtosed (too many people at the end)

It is a measure of flatness. It occurs when there are either too many people at the extremes of the scale or not enough people at the extremes.

.....

standard Normal Distribution

$$\text{If } X \sim N(\mu, \sigma)$$

$$\text{then } Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$

Here $Z$ = Standard Normal Distribution

* Skewness & Kurtosis can happen simultaneously

* ~~In~~ Even if population itself is not normal under the condition that $n \to \infty$ ( large sample is taken) even distributions like binonomial, poisson, etc. tend towards standard normal.

## Difference bet. parameter & statistics

A descriptive measure of population is known as a parameter, they are usually denoted by Greek letters eg. the mean of the population is denoted by '$\mu$' & standard deviation by '$\sigma$'. On the other hand, a descriptive measure of a sample is referred to as a statistic. They are usually denoted by Roman letters eg. mean of the sample is usually denoted by $\bar{x}$ & standard deviation of the sample is denoted by 's'

⇒ A study is as good as ~~a~~ the sample.

## Hypothesis

Ordinarily, when one talks about hypothesis, one simply means a mere assumption, assertion or supposition to be proved or disproved. For a researcher, hypothesis is a formal question that needs to be resolved. Technically, it is defined as a proposition or a set of propositions set forth as an explanation for the occurence of some specified group of phenomena either asserted merely as a provisional conjecture to guide some investigation or accepted as highly probable in the light of established facts.

Quite often, a research hypothesis is a predictive statements capable of being tested by scientific methods that relates an independent variable to some dependent variable, eg. students who receive counselling will show a greater increase in creativity than students not receiving counselling.

Hypothesis can be of 2 types:
(i) Simple → one statement
(ii) Compound → multiple things

* Characteristics of a good hypothesis:

(i) It should be a proposition that can be put to test to determine its validity. We should also be able to make other deductions from it which can lead to further research.

(ii) It should be clear, precise and reliable, clearly stating the relationship bet. the variables.

(iii) It should be limited in scope & specific. In fact, narrower, the better.

(iv) It should be consistent with a substantial body of established facts.

(v) We should be able to test the hypothesis within a reasonable time.

We have hypothesis classified as Null ($H_0$) & Alternative hypothesis ($H_1$ / $H_A$)

→ Null Hypothesis ($H_0$) :- There is no difference in performance bet. the different conditions you are studying.

→ Alternative Hypothesis ($H_1$) → There is a significant diff. bet. the conditions being studied.

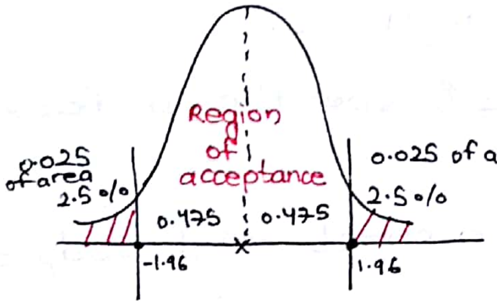
* Alternative Hypothesis is usually the one which one wishes to prove & the null hypothesis is the one which we wish to disprove. So, Null Hypothesis represents the statement, we want to reject & $H_1$ is all other possibilities.

There are only 2 outcomes to an inferential statistical analysis:

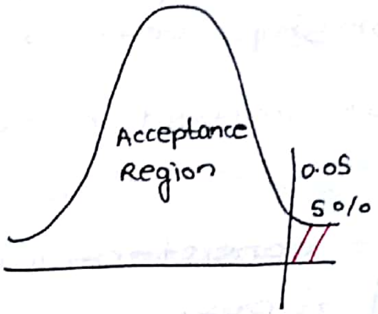
(1) The difference we find is due to some genuine & real effect. In this case, we reject $H_0$ that means we believe an effect truly happened.

(2) The difference, if at all, is only due to chance. (we fail to reject $H_0$, we never use the word 'accept' and we just say, that further research might be required)

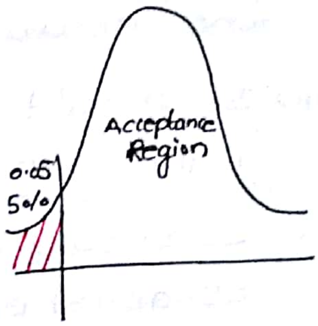
## Level of significance



Two-tailed rejection



Right tailed rejection



Left-tailed rejection

* Generally, we have 5 % level of significance

⇒ If value comes 2 → we reject Null Hypothesis

If it comes within -1.96 and 1.96 → we fail to reject null hypothesis.

* This is a very important concept in the context of hypothesis testing. It is always some %age, usually 5% which should be chosen with great care, thought & reason. In case, we take the significance level at 5 %, then this implies that $H_0$ will be rejected when the sampling result i.e. the observed evidence has a less than 0.05 probability of occuring if $H_0$ is true. When the probability is found to be below 0.05, it is described as statistically significant. Thus, the significance level is the maximum value of the probability of rejecting $H_0$ when it is true & is usually determined in advance before testing the hypothesis

⇒ Acceptance & rejection regions are in relation to $H_0$.

## *Decision Rule

Given a hypothesis $H_0$ & an alternative hypothesis $H_1$, we make a rule which is known as the decision rule acc. to which we reject $H_0$ or fail to reject it
eg. if $H_0$ is that a certain consignment is 'good' or not, we must decide the no. of items to be tested and the criteria for accepting or rejecting.

For statistical tests like z & t, level of significance helps us to decide the decision rule.

## *Type I error

Reject $H_0$ when it is actually true. Also known as a false +ve or a false hit. That means, wrongly infer the presence of something that is not.

## * Type II error

Fail to reject a false $H_0$. That means, there is an effect, cause or reason ~~tha~~ which has led to the difference but we fail to identify it. So, falsely infer the absence of something that actually is.

17-Jan-2019

The probability of type I error is usually determined in advance & is known as level of significance. We can control type I error by fixing the level of significance at a low level usually 5% or 2%. It is denoted by '$\alpha$'

Probability of type II error is usually denoted by the symbol '$\beta$'.

| Fact | Decision: Fail to reject $H_0$ | Decision: Reject $H_0$ |
|---|---|---|
| $H_0$ True | Correct Decision<br>$(1-\beta)$ | Type I error<br>$(\alpha)$ |
| $H_0$ False | Type II error<br>$(\beta)$ | Correct Decision<br>$(1-\alpha)$ |

$\alpha$ :- level of significance

$1-\beta$ :- Power of a test

$1-\alpha$ :- Confidence of a test

$P[\text{Type I error}] = \alpha$

$P[\text{Type II error}] = \beta$

$\boxed{\alpha + \beta \neq 1}$

With a fixed sample size 'N' when we try to reduce Type I error, the probability of commiting Type II error increases. So, there is a tradeoff bet. these 2 type of errors which means that the probability of making one type of error can only be reduced if we are willing to increase the prob. of making the other type of error. (For practical example refer to Signal Detection Theory).

To deal with the tradeoff in business situations, medical conditions, etc. decision makers decide the appropriate level of Type I error by examining the costs & penalties attached to both type of errors.

*Examples of Type I errors:

(i) Test that shows a patient to have a disease when in fact there is no disease.

(ii) Fire alarm going on indicating a fire when actually there is no fire.

(iii) ~~A~~ claim that a particular medical treatment cures a disease when it actually doesn't.

* Examples of Type II errors

(i) A test failing to detect the disease (which it was designed to detect) in a patient who really has the disease.

(ii) No fire alarm but actual fire

(iii) Failure of a clinical trial but the treatment actually works.

**Note:**

The statistical power of a study or a test which is also known as sensitivity determines how likely the study is to distinguish an actual effect from one of chance. It is the likelihood that the test is correctly rejecting the null hypothesis. So, power is 80 % means 80 % chance of the test having significant results.

## *Mean Variable

$$X \sim N(\mu, \sigma) \quad ; \; \sigma > 0$$

$$\begin{array}{cccc} X_1, & X_2 & \ldots\ldots & X_n \\ \downarrow & \downarrow & & \downarrow \\ N(\mu,\sigma) & N(\mu,\sigma) & & N(\mu,\sigma) \end{array}$$

IIDs
↳ Independent
↳ identically distributed

$\bar{X}$ = mean variable (sample mean)

$$= \frac{X_1 + X_2 + \ldots + X_n}{n}$$

( It is an unbiased estimator of $\mu$ i.e. average of $\bar{X} = \mu$ )

$$\boxed{\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)}$$

$$X \sim N(\mu, \sigma)$$

$$z = \frac{X - \mu}{\sigma}$$

$z \sim N(0, 1)$ [z-score]

$$\boxed{z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}}$$

If $X_1, X_2 \ldots\ldots X_n$ are IIDs (independent & identically distributed random variables) then $\bar{X} = X_1 + X_2 + \ldots X_n / n$ is called the mean ~~var~~ variable or the sample mean. It is a variable because it varies from sample to sample. If each $X_i$ has mean $\mu$ & standard deviation $\sigma$, we know that the mean variable will have a mean = $\mu$ & standard deviation = $\sigma/\sqrt{n}$. In particular, if each $X_i$ follows normal distribution with parameters $\mu$ & $\sigma$ then $\overline{X} \sim N(\mu, \sigma/\sqrt{n})$.

* Central Limit Theorem & its consequences

(i) If the distribution in the sample is normal i.e. each $X_i$ follows normal distribution, the $\overline{X}$ also follows normal distribution i.e. its frequency graph will be symmetrical, bell shaped, etc. irrespective of n.

(ii) If the distribution is not normal but the sample is large enough i.e. n goes towards infinity (for practical purposes $n \geq 30$) then the sampling distribution of $\overline{X}$ will again be normal.

In either of the 2 cases, we define $\boxed{z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}}$ which is known as the z-test

(we know $z \sim N(0,1)$)

## z-test

z-test is a parametric test which is used for comparing the mean of a sample to some hypothetical' mean for the population in case of ~~ea~~ either a large no. of samples or when variance is known. We generally assume sample to be normally distributed but it need not be so because of Central Limit theorem (if n is large).

It is also used for judging the significance of diff. bet. means of 2 independent samples in case of large samples.

Q) A sample of $\overset{(n)}{400}$ male students is found to have a mean height of $\underset{(\overline{x})}{67.47}$ inches. Can it be reasonably regarded as a sample from a large population with mean height $\overset{(\mu)}{67.39}$ inches & standard deviation 1.3". Test at 5 % level of significance.

Sol.

$n = 400$ $\qquad\qquad \mu = 67.39$

$\overline{x} = 64.47$ $\qquad\qquad \sigma = 1.3$

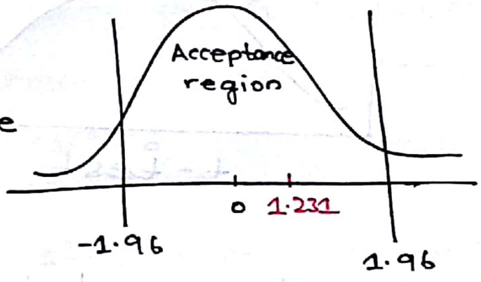
$H_0 : \mu = 67.39$

$H_1 : \mu \neq 67.39$

$$z = \frac{\overline{x} - \mu}{\sigma / \sqrt{n}}$$

$$= \frac{67.47 - 67.39}{1.3\sqrt{400}} = 1.231$$

5 % level of significance



$\therefore \mu = 67.39$

## t - test

$$\boxed{t = \frac{\overline{x} - \mu}{\sigma_s / \sqrt{n}}}$$

where $\sigma_s = \sqrt{\frac{\Sigma (x_i - \overline{x})^2}{n-1}}$ (sample standard deviation)

Degree of freedom $= n - 1$

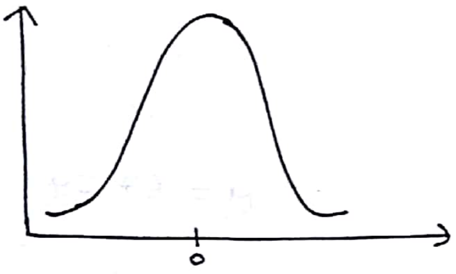
Early theoretical work on t-distribution was done by a scholar W.S. Gosett who published his papers under the pen name 'student'. consequently, the t-distribution is also known as 'the student's t-test'. It is mainly used in 2 scenarios:

(i) If the sample is small (especially if we do not know if sample follows normal) distribution.)

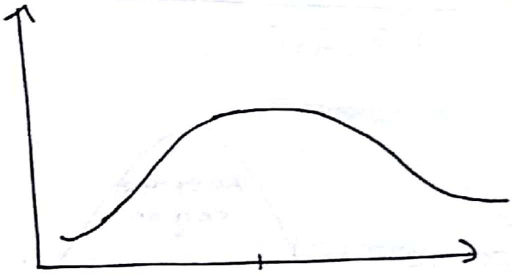
(ii) Population variance is not known.

## Chief characteristics:

(i) A t-distribution resembles the z-distribution in the sense that they are both symmetrical about a mean of zero value but the t-distribution is much more flatter & more spread out than the normal curve.



z-test



t-test

(ii) variance of t-distribution is equal to deg. of freedom (dof) / dof - 2

$$\text{variance} = \frac{dof}{dof - 2} = \frac{1}{1 - \frac{2}{dof}} = \frac{1}{1 - \frac{2}{n-1}}$$

(iii) As sample size increases, dof also increases & in this case as $n \rightarrow \infty$, t-distribution approximates or converges towards standard normal distribution.

## Applications

(i) To test the significance of mean
(ii) To test the significance of diff. bet. 2 means
(iii) To test the " " an observed correlation coeff.

(iv) To check the effectiveness of govt. schemes, policies, etc.

(v) In the field of Psychology, it can be used to compare the effect of various psychotherapies, memory techniques, utility of drugs, etc.

(vi) In the field of economics, it can be used to compare growth rates, per capita incomes, diff. in living standards, etc.

Q) What is degrees of freedom?

Ans= Degree of freedom of an estimate or a statistic is the no. of independent variables in the picture i.e. no. of variables which can be assigned values independently. For a single sample of size 'n' dof is 'n-1.' For 2 samples of order $n_1$ & $n_2$ combined dof is $n_1+n_2-2$

As degree of freedom increases, critical value falls.

Q) What is standard error?

Ans= The standard deviation of the mean variable i.e. $\sigma/\sqrt{n}$ is known as the standard error of the mean.

$$SE = \sigma/\sqrt{n} \quad ; \quad \sigma = \text{S.D. (population)}$$
$$n = \text{sample size}$$

So, clearly, it is affected by the sample size, bigger the size, smaller the standard error.

The second thing to affect it is the amount of variation in our sample.

If we have a lot of variation (more variance, therefore more S.D.), there will be more uncertainty about the population mean. So, it is a measure of dispersion of sample means around the population mean.

12-Jan-2019

# CORRELATION

There are 2 disciplines of scientific Psychology. One is correlational Psychology & other is experimental Psychology. Corr. Psychology is concerned with investigating the relationship bet. naturally occuring variables & studying individual diff. The experimental psychologist is not usually interested in individual diff. but rather with minimising of controlling these differences in order to show that some stimulus factor influences every individual's behaviour in some predictable way to some measurable deg. The correlationist observes variables & relates them whereas the experimentalist manipulates the variables & observes the outcome. The correlationist looks for ways in which people differ from each other & experimentalist looks for general laws that apply to everyone.

## Origin

£ Sir Francis Galton (1822 - 1911) discovered the main features of correlational analysis. He discounted the influence of env. (wrongly) and declared that 'genius' was a result of inheritance. Other things that he contributed include:

(i) He originated the drawing of scatter plots which led to modern day regression.

(ii) He discovered "Regression to the Mean" phenomenon. It refers to the tendency of a random variable that is highly distinct from the norm to return to the normal after repeated tests. He observed that taller than average parents had tall children but they tended to be not quite as tall as their parents. Similarly, shorter than average parents had shorter children but the weren't as short as their parents i.e. the children's height tended to drift back or regress to the mean of the population.

This regression to mean phenomenon can threaten a study's internal validity if not properly taken into account & is one of Galton's greatest discoveries.

(iii) Galton observed that a plot of the average of each column of his scatter plot yielded more or less a straight line. This in fact is a type of regression line.

* After Galton, Karl Pearson continued developing the idea eventually devising the modern formulae for calculating coefficient of correlation. He named it 'r' for regression in honour of Galton's discovery of regression ~~of~~ to the mean phenomenon.

* Concepts of Correlation

Central to Psychological Testing & assessment is finding out how somethings like traits, abilities or interests are related to other things such as behaviour or performance. A coefficient of correlation 'r' is a number that provides us with an index of the strength of the relationship bet. 2 variables ~~&~~ An understanding of the concept of correlation & ability to compute ~~the~~ coeff. of correlation is therefore central to the study of tests & measurement.

Simply stated correlation is an expression the degree & direction of correspondence between 2 things. Coeff. of correlation (r) expresses a linear relationship between 2 variables. It reflects the degree of variation & the extent to which the variables are related. The meaning of correlation coeff. (r) is interpreted by its sign (+ve or -ve) and by its magnitude. If 2 variables (X & Y) simultaneously increase or decrease then they are said to be +vely or directly correlated. In this case 'r' is +ve.

If increase in one variable leads to decrease in the other then they are said to be -vely correlated or inversely ~~co~~rrelated.

Also, greater the magnitude (absolute value), greater the degree of relatedness. The values of 'r' varies from -1 to 1 $(-1 \leq r \leq 1)$. $r = \pm 1$ indicates perfect correlation whereas $r = 0$ indicates uncorrelated which means no linear relationship exists between the variables.

~~Examples~~

Like mean & standard deviation, coeff. of correlation is also a descriptive statistic. The inferential analysis for correlations involves (i) getting the regression line which helps us to be predictive

~~(ii)~~

(ii) It involves determing if a particular correlation is significantly greater than or less than 0 i.e. in Correlational research

$H_0$ : value of $r = 0$
(i.e. no relationship exists)

$H_1$ : $r \neq 0$

Rejecting the null hypothesis means deciding that a significant relationship exists bet. the 2 variables

## *Coefficient of Determination ($r^2$)

It is easy to misinterpret the meaning of a particular Pearson's 'r'. If its equal to ·7, the relationship is relatively strong one but it does not mean that the relationship is true 70 % of the times ('r' does not mean %age)

A better sense of how to interpret the meaning of correlation is to use coeff. of determination i.e. $r^2$. Since, it is found by squaring 'r', therefore, it is always a +ve number regardless of whether correlation is +ve or -ve. Technically, it is defined as the portion of variability in one of the variables being used in the correlation that can be accounted for by the variability in the second variable.

Eg. If $r_1 = 0.5$ $r_1^2 = 0.25$ (25 %)

$r_2 = 0.7$ $r_2^2 = 0.49$ (49 %)

Although, we may be tempted to think that the relationships are not all that diff. in the eg. i.e. 0.7 is not that much bigger than 0.5, the reality is that the amount of variability bet. the 2 data is almost twice.

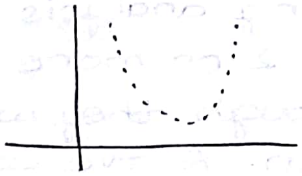
16 - Jan - 2019

## * Limitations of correlation

↳(i) It does not imply cause. With its ability to predict, correlational research is a powerful tool. However, finding a correlation does not allow the conclusion that one of the variables is causing the other to occur. Unfortunately, a failure to appreciate this fundamental rule makes correlational research the method least understood by not uncommon for news reports to describe correlational research & for the narrative of the report to suggest to a non critical reader that one variable in the relationship is the cause to the other.

In an experimental study with a manipulated independent variable, cause & effect conclusions can be drawn with some deg. of confidence. The independent variable of interest is manipulated & if all else is held constant, the results can be attributed directly to the independent variable. With correlational research, the "all is held constant" feature is missing and this lack of control makes it impossible to conclude anything about cause & effect (Directionality problem not answered).

(ii) Correlation is a measure of linear relationship only. Pearson's correlation coefficient fails to identify the nature of a non- linear relationship.



In case the scatter plot is like above, we get the value of 'r' almost equal to zero establishing no correlation but that only means there is no linear relationship, they could be another curve of higher degree through which the points pass.

(iii) The misleading influence of a third variable, an underlying third variable say 'z' to which both X & Y, may be related can often may be detected & can infact be the reason for correlation bet. X & Y ('x' & 'y' → the values the variable takes [small letters])

Galton & Pearson discounted the effects of env. as the factor underlying the relationship bet. parent's & children's intellectual abilities. +ve correlation bet. parent's & children's intelligence helped them conclude that heredity was the prime cause of diff. in level of intelligence. A probable third variable is the fact that intelligent parents were able to provide a stimulating env. to their children.

(iv) Spurious correlation of a part with the whole.

If we try to correlate 2 variables, one of which forms part of other, we shall find an arithmetic relation which may easily lead to a significant value of 'r' but it does not help us to analyse the situation eg. if we find correlation bet. 'X' & 'X/Y'

(v) Combination of unlike populations

Another way we may get a wrong analysis of correlation is when data from 2 or more separate groups are combined as though they were a single homogeneous population. A +ve correlation in one as compared to a -ve correlation in the other may lead to a negligible correlation when the groups are combined.

(vi) Inference to an unsampled population

When correlation is used for predictive purposes (regression), it is statistically dangerous to infer something about a wider population than that covered by the experiment.

(vii) Illusory correlation

eg. Women are bad drivers

(there seems to be a correlation but it is false)

## • The need for Correlational Research

↳ There are many pitfalls that exist when trying to interpret correlations, still there is a need for correlational research due to both practical & ethical considerations. On fairly obvious practical grounds, some imp. research would not be feasible or even possible as a pure experimental study. Studying gender differences in behaviour, diff. bet. age groups or diff. bet. personality type are major research areas in which it is not possible to randomly assign (which is the key factor of experimental study)

A study looking for a correlation bet. the severity of clinical depression & task perseverance has value in its own right even if the subjects cannot be randomly assigned to high depression & low depression conditions.

Furthermore, some research is conducted with only prediction as the major goal eg. finding a correlation bet. certain personality attributes & job success allows the company managers to predict who will succeed in the job while remaining unconcerned about why the relationship exists.

Also on rather clear ethical grounds, there are some studies that simply cannot be done as experiments with manipulated variables eg. a french Physiologist Paul Broca discovered the brain's speech centre by noting a relationship bet. certain types of speech disorders & the extent of brain damage discovered on post mortem examination. Experimental evidence that the disorder was caused by the brain damage would require randomly assigning people to a brain damaged group which will have portions of their brains pulverised or to a safer control group. Human volunteers are difficult to find & this is one reason why animals are used as subjects in exp. research investigating the relationship bet. brain & behaviour.

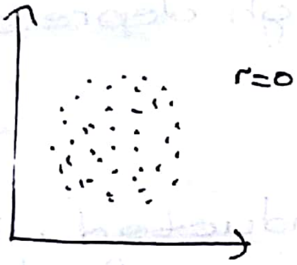
Research using correlational procedures can be found in all areas of psychology esp. prevalent in:

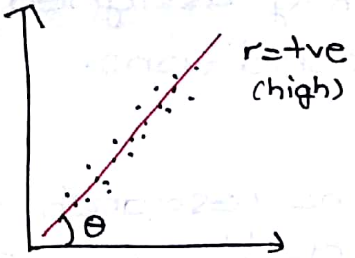
(i) Research concerning dev. of psychological tests to check reliability & validity.

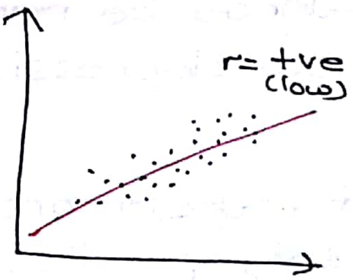
(ii) Research in personality types

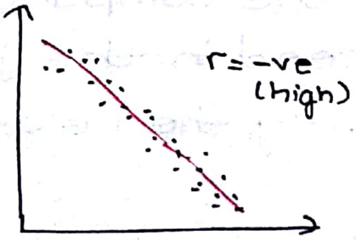
(iii) Research in nature vs nurture (study on twins)

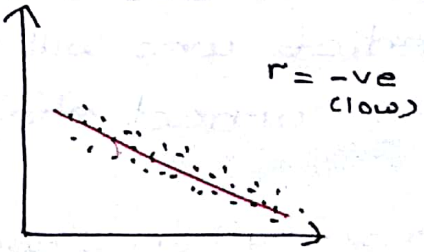
## Regression











- θ= 45° (Perfect correlation) {more or less θ will reduce correlation}

| |
|---|
| r= 1 |
| r= -1 |

Regression analysis helps us to make prediction. It should be performed only when strong correlations exist. If you know that a statistically significant correlation exists bet. 2 variables (by calculating the value of 'r') then knowing a score on one of the variables enables you to predict a score on the other. Here, we draw a regression line through a scatter plot which is a line of "best fit" i.e. it provides the best possible way of summarising the points with the help of a line. More precisely, if you look at the absolute values of the shortest distances bet. each point & the line, those distances would be at minimum. If the regression line is $Y = mX + c$ then 'm' is the slope which can be calculated if we know:

(i) the correlation coefficient (r).

(ii) the standard deviations for the variables being correlated. $(\sigma_x, \sigma_y)$

c = y intercept which can be calculated if we know

(i) value of m

(ii) mean scores for the variables being correlated $(\mu_x, \mu_y)$

$\boxed{Y = mX + C}$

This is called the regression of Y on X, 'Y' is the dependent variable also known as the criteria variable where as X is independent variable also know as predictor variable. We can also have regression of X on Y.

In order to predict with confidence, the correlation must be significantly away from 0, higher the correlation, closer the points on the scatter ~~points~~ plot will be to the regression line &

more confident you can be in the prediction you make eg. A report about risk factors for Heart Attack will be describing a study in which a significant correlation bet. say smoking & heart disease allows the prediction that heavy smokers will be more likely to develop heart problems ~~that~~ than non-smokers i.e. smoking predicts heart disease.

23 - Jan - 2019

## Z - score

$X \sim N(\mu, \sigma)$

Then, $Z = \dfrac{X - \mu}{\sigma}$

$Z \sim N(0, 1)$

Q) Evaluate the importance of deviation IQ in the measurement of intelligence.

Ans = Deviation IQ is calculated as how much someone deviates from the average IQ of 100.

Introduction of mean scores to calculate IQ led to a few complications because researchers found that they were dealing with much more info. by introducing mean scores across a huge age range on an intelligence test. eg. 103.2 is the average IQ test score for 32 year olds, 102.2 is the average IQ test score for 33 year olds etc. This led to issues for allowing comparisions.

Suppose, for example, a recruiter has 2 candidates. Person A aged 20 years who scored 105 on the intelligence test (115 being the avg. intelligence test score among 20 year olds) & person B aged 25 years scored a 100 on the intelligence test. (105 being the avg. intelligence test score for 25 year olds). Which candidate should be hired?

* Wechsler realised that ~~he~~ we had to standardise the scoring of IQ, the aim was to find a way of comparing all these diff. scores & still provide a standardised scoring system. ~~A~~ For this he used Normal distribution curve.

Researchers have found that many of the variables measuring human attitude & behaviour including intelligence follow normal distribution curve. Wechsler decided to use 100 as an avg. IQ based on the original IQ calculations & standard Deviation 15. Using the normalising ~~prod~~ procedure, we can calculate z-score and analyse it. The usual interpretation is:

(i) A score that is no more than 1 standard Deviation away from 100 can be interpreted as a normal score (i.e. a score between 85 and 115)

(ii) A score that is between 1 & 2 standard deviations away from 100 can be interpreted as low (70-85) & high (115-130).

(iii) A score that is more than 2 standard deviations away from 100 (lower than 70 or higher than 130) can be interpreted as very low or very high.

Therefore, he provided a benchmark by which all people of all ages could be compared in IQ by using standard deviation.

# Multivariate Regression

A bivariate approach investigates the relationships bet: any 2 variables. A multivariate approach on the other hand examines the relationship among more than 2 variables.

→Multivariate Regression solves the problem of having more than one predictor of some outcome. There is one criteria variable (dependent $y$) and a minimum of 2 predictor variables (independent $x_1, x_2$).

→The analysis enables you to determine not just the fact that these 2 or more variables combine to predict some criteria but also the relative strengths of the predictors known as weights of coefficient.

$$\boxed{Y = a_1X_1 + a_2X_2 + \ldots\ldots + a_nX_n + b}$$

The advantage of a multiple regression analysis is that when the influences of severable predictor variables ~~predi~~ are combined, prediction improves far beyond the single regression case. This is especially true if the predictors are not highly correlated with each other.

A study predicting empathy from 2 aspects of early childhood experience. Empathetic adults tended to have experienced distressing childhood events that could have made them more sensitive to trauma of others. The severity of the trauma was weighted more heavily as a predictor than the no. of traumatic events.

# ANOVA

(f- test)

ANOVA is ~~an~~ the analysis of variance. Sir Ronald Fischer who is considered to be 'Father of Modern Statistics is the one who introduced variance & developed an elaborate theory of ANOVA explaining its usefulness in practical field. It is basically a collection of statistical models & their associated estimation procedures used to analyse the diff. among 'group means' in a sample. It is similar in application to techniques such a z-test & t-test in the sense that it is also used to compare means & relative variance between them. However, ANOVA is best applied where more than 2 populations or samples are meant to be compared. The essence of ANOVA is that total amount of variation in a set of data is broken down into 2 types:

(i) The amount which can be attributed to chance.

(ii) The amount which can be attributed to specified causes.

There may be variation between samples & also within the sample items. ANOVA consists in splitting the variance for analytical purposes. Through this technique, one can explain whether the various varieties of seeds or fertilisers or soils differ significantly, so that a policy decision could be taken accordingly concerning a particular variety in the context of agricultural research. Similarly, mileage of diff. automobiles or various types of drugs manufactured for curing a specific disease may be studied & judged to be significant or not, through the application of ANOVA. In other words ANOVA is used to test the significance of differences among the means of a no. of diff. groups. ~~Does~~ ~~does~~ Thus, through this technique one can in general investigate any no. of factors which are hypothesised or said to influence the ~~def~~ ~~dede~~ dependent variable. One may as well investigate the diff. among various categories within each of these factors.

If we take only one factor & investigate the diff. amongst its various categories, we are said to use One way ANOVA.

In case we investigate 2 factors at the same time, then we use Two Way ANOVA. In a 2 or more way ANOVA, the interaction i.e. the inter-relationship bet. 2 independend variables or factors (if any) can also be studied for better decisions.

eg. If we want to compare whether or not the mean output of 3 workers is same based on their working hours then we are using 1 way ANOVA.

If we do the same study based on working hours & working conditions then we are using a 2 way ANOVA.

25-Jan-2019

In ANOVA, we measure an outcome variable, also called the ~~a~~ dependent variable. This outcome variable must be measured on a continuous scale. It depends on one or more predictor variables which may be continuous or categorical. Usually in ANOVA predictor variables are categorical. In this case, they are known as factors or independent variables. These might be variables that we manipulate like a treatment or they might be variables that we simply measure eg. height, weight, gender, etc.

Whenever we measure anything, we find differences. 'Measuring' people on our outcome measure is no exception. These diff. can happen because of 2 reasons.

(i) Because of the effect of the predictor variable.

(ii) Because of other reasons (including chance factor).

We want to know 2 things:

(i) How much of the variance bet. different groups is due to the predictor variable.

(ii) Whether this proportion of variance is statistically significant i.e. is it larger than what we would expect by chance if the null hypothesis was true.

Example:

→Wright (2005)

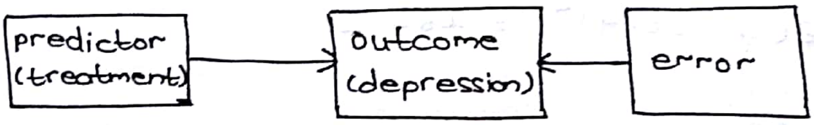
A study was carried out to compare the effectiveness of 2 diff. treatments for chronic fatigue syndrome in children. First treatment known as 'Pacing' encourages the child to exercise to the point of tolerance but to avoid over-exertion & resting when necessary but trying to avoid total rest. Second treatment was called 'Active Rehabilitation Programme'. Here it was similar to the pacing programme but it also included cognitive behavioural therapy to try to change the child's attributions about their condition and to help them with active coping strategies. One of the variables assessed was depression.

People score differently on a depression measure for a very wide range of reasons but we want to know whether one of those reasons was the treatment that was given. The diff. bet. the scores on depression are represented by the variance of depression score.

eg.

| z-score | | proportion scoring lower | | % |
|---|---|---|---|---|
| 3.5 | → | 0.9998 | → | 100 % |
| 2.0 | → | 0.9772 | → | 98 % |
| 1.5 | → | 0.9332 | → | 93 % |
| 1 | → | 0.8415 | → | 84 % |
| 0 | → | 0.5000 | → | 50 % |
| -1 | → | 0.1587 | → | 16 % |
| -1.5 | → | 0.0668 | → | 7 % |
| -2.0 | → | 0.0228 | → | 2 % |
| -3.5 | → | 0.0002 | → | 0 % |

## Compartmental Diagram



Some of the variance ~~is~~ (diff. bet. people) is due to the treatment whereas some could be due to other factors maybe including:

(i) How depressed the children were to start with.

(ii) Uncertainty in the measurement tech.

(iii) Day to day changes in mood amongst children.

In the ANOVA technique into 3 categories:

(i) Total Variance

(ii) Variance due to treatment (the diff. bet. 2 groups)

(iii) Variance due to error (difference within each group)

Variance is conceptualized as sum of squared deviations from the mean which is shortened as sum of squares (SS) $\Sigma(x-\overline{x})^2$

*(Avg. hours of exercise) per month

| | Group A (Pacing) (X) | Group B (ARP) (Y) |
|---|---|---|
| | 14 | 11 |
| | 16 | 11 |
| | 9 | 9 |
| | 18 | 18 |
| | 13 | 3 |
| Mean | 14 | 10.4 |

## Three kind of SS

(i) $SS_{TOTAL}$

Total Average $= \overline{X} = 12.2$

$\Sigma(X-\overline{X})^2 = 193.6$ [including both groups]

(ii) $SS_{WITHIN}$

$$\Sigma(x-14)^2 + \Sigma(y-10.4)^2 = 161.2$$

(iii) $SS_{BETWEEN} + SS_{WITHIN} = SS_{TOTAL}$

$$SS_{BETWEEN} = SS_{TOTAL} - SS_{WITHIN}$$
$$= 193.6 - 161.2$$
$$= 32.4$$

We want to know how large the effect of the treatment has been. This is same as asking-"what proportion of the total variance the treatment effect has been responsible for?"

we define $\boxed{\text{Effect Size } (R^2 \text{ or } \eta^2) = \dfrac{SS_{BETWEEN}}{SS_{TOTAL}}} \approx \left(\dfrac{Var_{BETWEEN}}{Var_{TOTAL}}\right)$

$$\frac{32.4}{193.6} = 0.167$$

The proportion of variance due to the treatment is 16.7% and therefore 83.3 % is due to error (error includes chance)

Now, let us calculate statistical significance.

(Inferential part of ANOVA is f-test)

## f-test

**Step 1: Calculate the degrees of freedom**

$df_{total}$, $df_{between}$, $df_{within}$

$$df_{total} = df_{between} + df_{within}$$

$$df_{total} = 10 - 1 = 9 \qquad [n-1 = df_{total}]$$

~~$df_{total}$~~

$$df_{between} = \text{No. of groups} - 1 = 2 - 1 = 1$$

$$df_{within} = 8$$

**Step 2: Calculate MS (Mean Squares)**

$$MS_{between} = \frac{SS_{between}}{df_{between}} = \frac{32.4}{1} = 1$$

$$MS_{within} = \frac{SS_{within}}{df_{within}} = \frac{161.2}{8} = 20.15$$

**Step 3 : Calculate f**

('f' is a test statistic for ANOVA. Also known as 'f-ratio')

$$\boxed{F = \frac{MS_{between}}{MS_{within}}} = \frac{32.4}{20.15} = 1.68$$

Null Hypothesis is that the avg. performance in the 2 groups is same.

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

Step 4:

$$p\text{-value} = F(1.68, df_{between}, df_{within})$$

$$= F(1.68, 1, 8)$$

$$= 0.24$$

Since, 0.24 is greater than 0.05, we can't reject the Null Hypothesis. that means that this study has not provided evidence that there is any difference bet. the 2 methods.

Note: If we analyse the same data using t-test, the critical value or p-value comes out to be exactly 0.24

we shouldn't be surprised by this because when we have 2 groups f-test & t-test are identical. In general

$$\boxed{t^2 = f}$$

Assumptions in ANOVA

(i) Normality within each group.

Outcome may not be normal but data within each group is assumed to be normally distributed

(ii) Homogenity of variance

As with the t-test, we assume that the S.D. within each group is approx. equal which is often referred to as homogenity of variance.

We don't need to worry about this assumption if we have approx. equal no. of people in each group.

There are 2 main reasons for this:

a) We use the variance to estimate the standard 40 error which obviously we want as low as possible.

b) If we can assume homogenity of variance, we can pool the variance from each group thus making the calculation easier.

(iii) Additivity of Effects

The basic model of ANOVA states that a given observation or score is a sum of certain components each due to the effect of a particular identifiable source of variation. In most cases, this assumption is generally met.

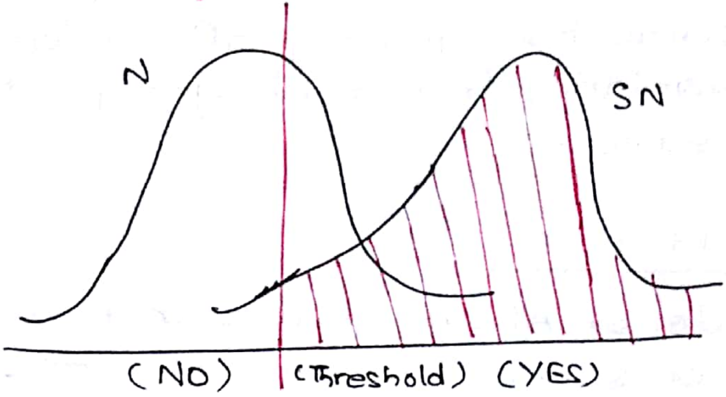
(iv) Random Sampling

The sampling within the various sets should be random. It usually means that observations are mutually independent & with equal chance of selection.

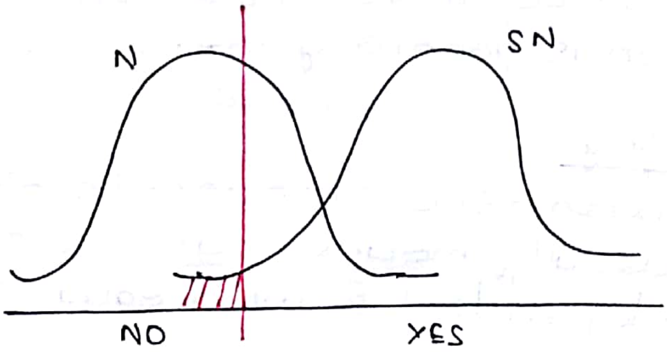
29-Jan-2019

# Signal Detection Theory

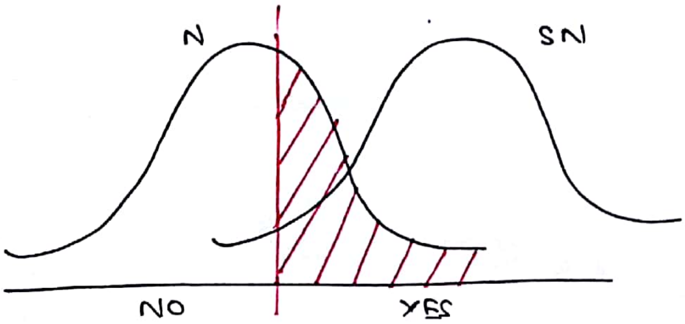
## I. HIT


N
SN
(NO)
(Threshold)
(YES)

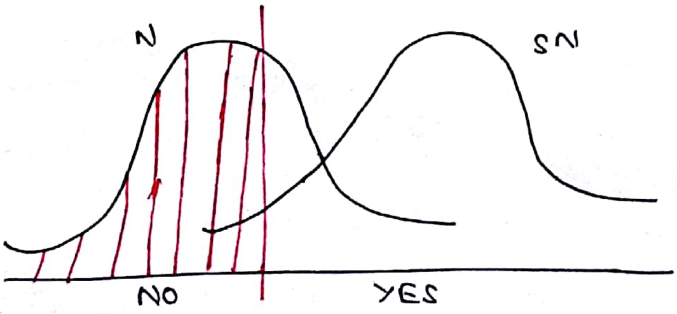
## II. MISS



## III. FALSE ALARM



## IV. CORRECT REJECTION



* ~~check~~ See only one graph at a time(while drawing)

* Low Threshold:- Both Hit rate & False Alarm rate increase.

↳ with high criterion, Hit rate & False Alarm rate reduces.

31-Jan-2019

# Psychophysics

## * Threshold

* Psychophysics as the name suggests is the study of relationship bet. magnitude of sensation & the physical intensity of the stimulus that evoked the sensation Thus, the area of psychophysics is the measurement of perception. One of the major areas it deals with is that of detection. Detection refers to the perception of presence of the stimulus. The problem of detection is concerned with the intensity of the stimulus necessary for its presence being sensed. The minimal value of stimulus which is perceptible is called the 'Thereshold'. (also called 'Limen') [German]

(Difference bet. minimum & minimal — find)

There are 2 types of threshold:

(i) Absolute

(ii) Differential

* Absolute threshold refers to the minimum amount of physical stimulus that produces a sensation.
eg. clicking of a clock in a room is audible with gaps. i.e. sometimes we hear it & sometimes we don't. It means that the threshold keeps on shifting due to momentary variations in the sensibility of receptors & makes the measurement of threshold difficult. Hence, we cannot conceive it as an abrupt change from a stimulus value that gets no response to a value that gets the response.

Therefore, statistically we define threshold as the minimal amount of physical stimulus that produces a sensation in 50% of the trials.

* Differential Threshold is defined as the minimal stimulus change which is required for the experience of different sensation. In other words, it is

the minimum difference in intensity between 2 stimuli that one can detect.

(Differential threshold is more imp. than Absolute Threshold)

Note:- A low threshold means that only a low intensity of the stimulus is required to detect the stimulus i.e. we are sensitive to that stimuli. Lower the threshold, higher the sensitivity. Thus, index of sensitivity is threshold.

## * Measurement of Threshold

### (i) Method of Limits

In this method, the stimulus intensity is decreased or increased in equal steps and in a serial order. (either ascending or descending → not both.). It is also known as a method of minimal changes or method of serial exploration eg. to measure absolute threshold of audible stimulus intensity, we first fix the range, say from 14 units to 23 units when it is known that the subject usually hears 18 scale units. The steps are to be kept equispaced and minimal. The stimuli are presented in serial order to which the subject's response is 'yes' or 'no' (to the presence of the signal). We present the stimuli in 2 modes. In one mode, the experimenter starts with a stimulus value well above the perceptible value & goes on decreasing the intensity in equal steps on successive trials till there is a change in the response from 'yes' to 'no.' This mode of presentation is known as descending series.

Opposite of this mode is ascending series.

The ascending & descending trials are presented alternately. Also, as a precaution to reduce the error the starting point is made to differ for different trials.

- Avg. of the threshold values obtained gives the value of absolute threshold.
- Separate averages for ↑ & ↓ trials gives the range of absolute threshold.

* Type of errors in measuring Threshold

(i) Error of habituation

↳ It refers to the tendency of the subject to continue the same response even when a distinct change in the stimulus has ~~occurred~~ occurred.

↳ Such a tendency results in an inflated threshold in an ↑ series & low threshold in ↓ series.

(ii) Error of anticipation

↳ It refers to the tendency of a subject to expect a change when in reality no distinct change has taken place.

↳ The subject feels he has said 'yes' so many times that now he should say 'no'.

Note:- Usually both these errors work in opposite directions & hence cannot be present in the same subject.

(iii) Time error

↳ It refers to overestimation or underestimation in judgements due to 'time factor' presentation of stimuli. The stimulus presented first is generally judged to be less intense. This error exists almost in all psychophysical measurements where presentation of 2 stimuli differ in time but has not been scientifically & satisfactorily explained.

(iv) Space error

↳ The tendency of being influenced by the stimulus being placed right or left of the subject & maybe by its position up or down results in space error.

(v) Movement Error

↳ The movement of the stimulus esp. when the subject adjusts the intensity of the stimulus himself, affects his judgement. This may lead to a systematic error of movement depending upon the direction i.e. inward or outward with respect to the subject.

2-Feb-2019

(ii) Method of constants

It differs from the method of limits in the mode of presentation of the stimuli. Here the mode of presentation is random. A stimulus of high intensity may be followed by a very low intensity stimulus. Thus, method of constants is free from errors of habituation & anticipation. The experimenter presents each of the fixed or constant set of stimuli for a constant no. of times. Hence, the name. The frequency of response for each stimulus is what the experimenter counts that is why, it is also known as the method of frequency. The frequency is converted to proportion by dividing with the total no. of trials.

One of the biggest advantage here is that we do not have to present the stimuli in equispaced fashion.

This method is applied to a very wide variety of problems like those of attention, memory & mental tests. This method also has a high degree of accuracy.

- Disadvantage:- It is not very economical from the point of view of time & labour.

Also, a great amount of precaution has to be taken with the help of pre-test trials in which the experimenter makes a rough estimate of the threshold.

(iii) Method of average error

(also known as method of adjustment)

↳ In the other two classical psychophysical methods, the control over changes in the stimuli was entirely in the hands of the experimenter. However, in this method, the subject himself manipulates the stimulus. Thus, the subject has to adjust the variable stimulus in such a way that he experiences a particular sensation. eg. the subject may be asked to adjust the knob in such a way that he can hear the sound. This is generally done from a higher intensity to just audible sound in half the trials and from imperceptible to just audible in another half. The thresholds found by this method are usually not very reliable and they are more of rapid approximations. This method is not generally used for determining threshold but to determine equivalents.

(i) It is economical in the sense that in each judgement, we get a measurement. saves time & labour as compared to the other 2 methods.

(ii) However, controlling the intensity may have a bad impact on the subject's judgement esp. when the handling of the apparatus is complicated

16-Feb-2019

## Research ~~Experimental~~ Designs

## (Experimental & Ex-post Facto)

Ex post facto research is ideal for conducting social research when it is not possible or acceptable to manipulate the characteristics of human participants. It is a substitute for true experimental research & can be used to test hypothesis about cause & effect or correlational relationship where it is not practical or ethical to apply a true experimental design. It literally means 'from what is done afterwards'. It can be viewed as experimental research in reverse. Instead of taking groups that are equivalent & subjecting them to diff. treatments to determine diff. in dependent variables, an ex-post facto experiment begins with groups that are already different in some respect & searches in retrospect for factors that brought about those differences.

What is Research Design?

⇒ It is the arrangement of conditions for collection & analysis of data in a manner that aims to combine relevance to the research purpose with economy in procedure. It is a conceptual structure within which research is conducted.

Need

(i) It constitutes the blueprint for the collection, measurement & analysis of data.

(ii) Facilitates smooth sailing of various research operations.

(iii) Helps to look for flaws & inadequacies. If research design is not carefully thought, it can give rise to misleading results or futile effort.

- Features

(i) If the research happens to be exploratory or formulative one where the major emphasis is on discovery of ideas & insights. The RD must be flexible enough to permit the consideration of many diff. aspects of the phenomena.

(ii) When the purpose of study is accurate description of a situation or of an association bet. variables, accuracy becomes a major consideration & RD which minimises the bias & maximises the reliability of evidence is considered a good design.

- Difference bet. experimental & non-experimental research

- Research in which the independent variable is manipulated is known as experimental hypothesis testing.

- In non-experimental → IV not manipulated

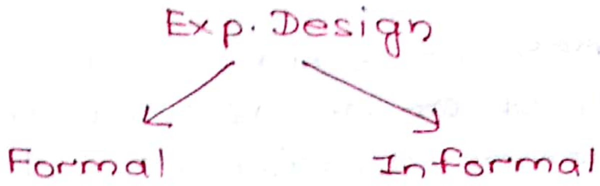
→ eg. suppose a researcher wants to study whether intelligence affects reading ability for a group of students

For this purpose, he randomly selects 50 students & tests their intelligence & reading ability by calculating the coeff. of correlation bet. the 2 set of scores. This is non experimental research because intelligence was not manipulated.

Suppose 50 students were divided into 2 random groups & one of the group was given special training then to check the effectiveness of training on intelligence, would be performing an experimental research.

The 3 basic principles of experimental design are:

(i) Principle of Replication
(ii) " " Randomization
(iii) " " local control



- Formal :- statistical procedures for analysis can be carried out.

- Informal

a) Before & After without control Design
b) After only with control Design
c) Before & After with Control Design

- Formal

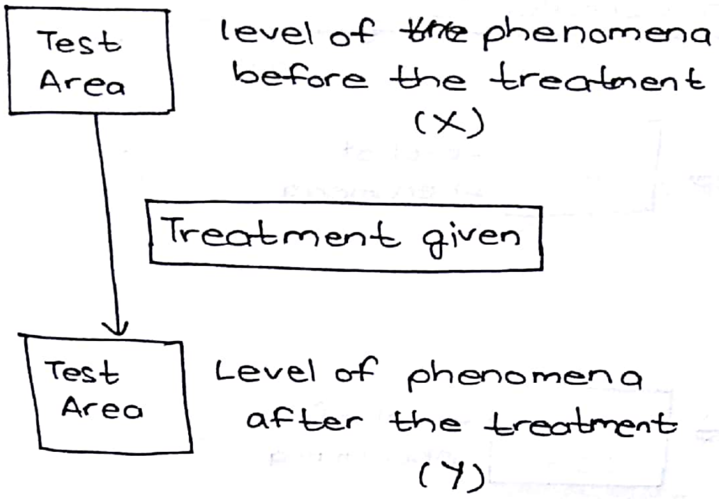
a) completely Randomized Design (CR Design)
b) Randomized Block Design (RB Design)
c) Latin Square Design (LS Design)
d) Factorial Design

* Informal

a) Before & After without control Design

A single test group or area is selected & the dependent variable is measured before the introduction of the treatment. The treatment is then introduced & DV is measured again.

Effect of the treatment is equal to level of the phenomena after the treatment − level of the phenomena before the treatment.

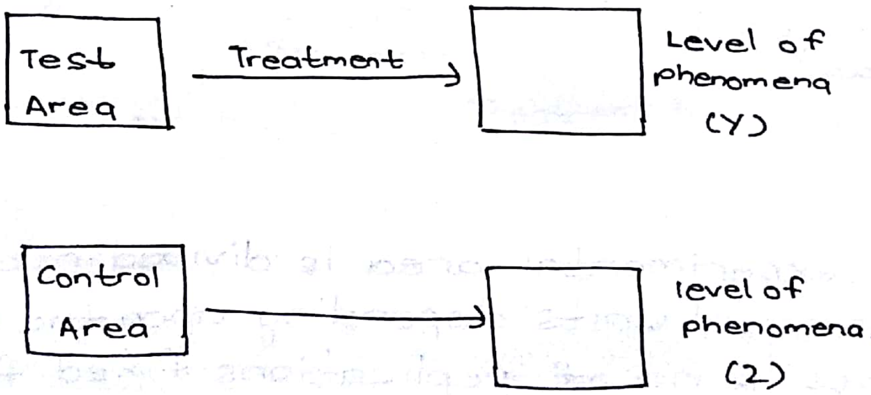

Test Area
level of ~~the~~ phenomena before the treatment (X)
Treatment given
Test Area
Level of phenomena after the treatment (Y)

Effect of Treatment = Y − X

Disadvantage

With the passage of time, extraneous variables may come to play a major role.

b) After only with control Design


Test Area
Treatment
Level of phenomena (Y)
Control Area
level of phenomena (Z)

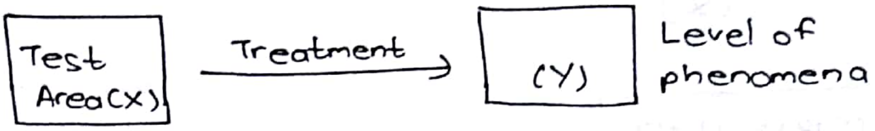
Treatment Effect = Y − Z

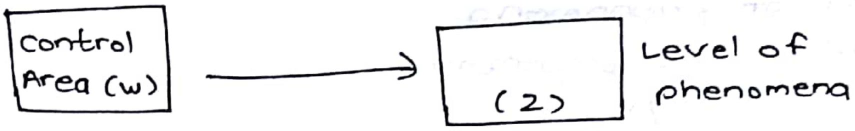
- In this design, 2 groups are selected & treatment is given only in one (Test area), the other group is called control area. The DV is then measured in both the areas at the same time.

**Disadvantage**

It is assumed that the 2 areas are identical w.r.t. their behavior towards the phenomena being considered.

c) **Before & After with control Design**





Treatment Effect = (Y−X) − (Z−W)

It is more superior than the previous 2, since it avoids extraneous variation resulting both from passage of time and from non comparability of the test & control areas. But at times, due to lack of historical data, it may not be feasible to formulate this design.

(Refer the notes)

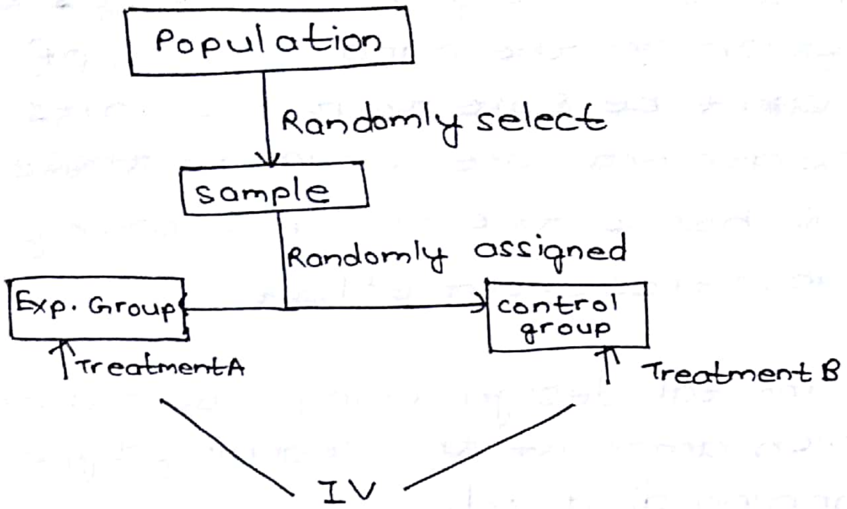
* **Formal**

a) **CR Design**

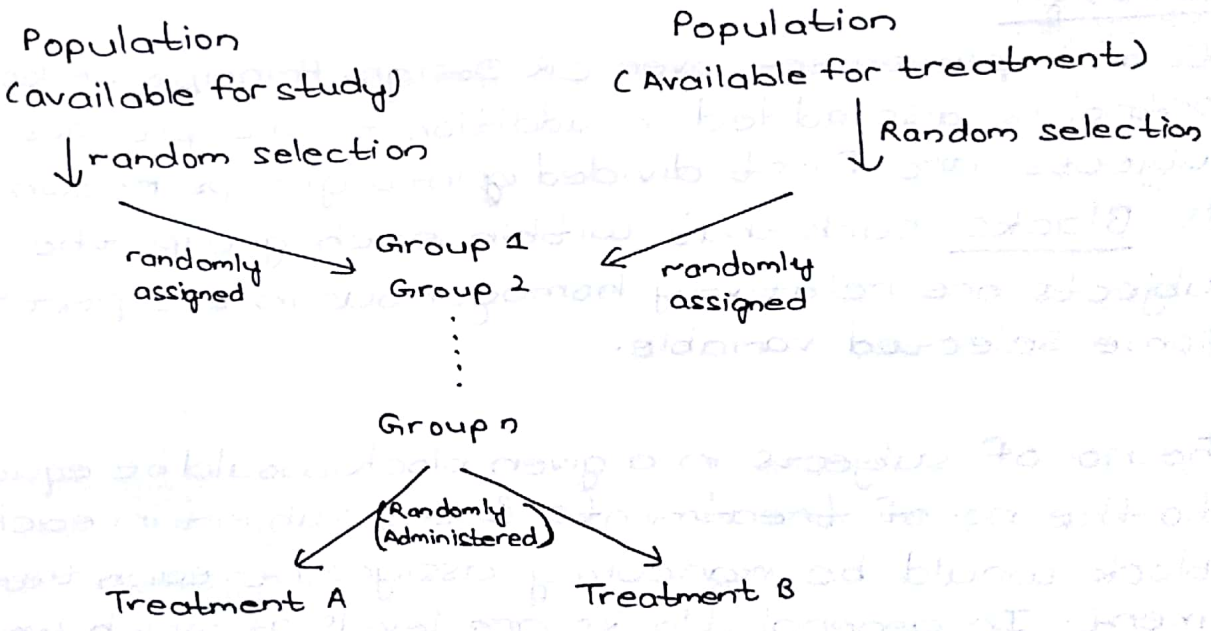
The whole experimental area is divided into a no. of experimental units depending upon the no. of treatments & no. of replications fixed for each treatment. After that the treatments are alloted to the units entirely by chance & DV is measured.

There are 2 main types of CR Design

(i) **Two group simple randomized design**


Population
Randomly select
sample
Randomly assigned
Exp. Group
control group
Treatment A
Treatment B
IV

## (ii) Random Replication Design


Population
(available for study)
random selection
Population
(Available for treatment)
Random selection
randomly assigned
Group 1
Group 2
randomly assigned
Group n
(Randomly Administered)
Treatment A
Treatment B

~~RB Design~~

## Merits

(i) A no. of treatments & replicates can be used.

(ii) The no. of replicates can be varied from treatment to treatment.

(iii) This design is esp. useful in small experiments where sample is homogenous.

(iv) Relative loss of info. due to missing data is smaller than any other design.

The main objection to this design is accuracy. since thereis no restriction on the randomisation of treatments, we can't be sure that the units receiving one treatment are similar to those receiving other & hence the variation among the units enter into treatment effect.

Q) Devise an experimental design which would allow you to test which amongst the learning types is best for learning disabled.

Ans= Behavioral, Cognitive, Psychodynamic

b) RB Design

It's an improvement over CR Design. Principle of local control is also added in addition to the previous 2. Subjects are first divided into groups known as Blocks such that within each group, the subjects are relatively homogenous in respect to some selected variable.

The no. of subjects in a given block would be equal to the no. of treatments & one subject in each block would be randomly assigned to each treatment. In general, blocks are levels at which we hold the extraneous factors fixed so that it's contribution to the total variability to the data can be measured.

2-Mar-2019

(iii) Latin Square Design

This design tries to control variation in 2 directions simultaneously i.e. it uses 2 blocking factors.

The blocking factors are represented through rows and columns. This design requires the no. of rows, no. of colums & treatments to be equal. Most commonly we have a 4x4 or a 9x9 design

## (iv) Factorial Design

Here 2 or more independent variables are there. Every level of 1 factor is studied against every level of another factor. They are esp. important in economic & social phenomenon where usually a large no. of factors affect a particular problem.

(goto page 57)

continued on
Page 57

2-Mar-2019

# FACTOR ANALYSIS

occupation
stay
study
income

→ combined together to make a Factor ($F_1$)

⇒ Correlated variable can also be used here.

* Factor → combination of variables which are highly correlated (are clubbed together)

## *Example I

$V_1$ → Paragraph comprehension

$V_2$ → vocabulary

$V_3$ → Arithmetic Problem Solving

$V_4$ → Statistical Reasoning

5 people → A, B, C, D, E

$$\begin{array}{c|ccccc} & A & B & C & D & E \\ \hline V_1 & & & & & \\ V_2 & & & & & \\ V_3 & & & & & \\ V_4 & & & & & \end{array}$$

Data Matrix

Correlational Matrix

$$\begin{array}{c|cccc} & V_1 & V_2 & V_3 & V_4 \\ \hline V_1 & 1 & a & b & c \\ V_2 & a & 1 & d & e \\ V_3 & b & d & 1 & f \\ V_4 & c & e & f & 1 \end{array}$$

a = 0.49
b = 0.24
c = 0.07
d = 0.24
e = 0.20
f = 0.48

## * Factor Matrix

F1 : Verbal Ability

F2 : Mathematical Ability

$$\begin{array}{c|cc} & F1 & F2 \\ \hline V_1 & 0.7 & 0.1 \\ V_2 & 0.7 & 0 \\ V_3 & 0.1 & 0.7 \\ V_4 & 0.6 & 0.6 \end{array}$$

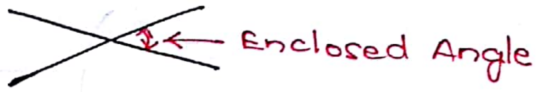
* John Carroll (1993) → "Three Stratum Theory"

*Factor loading → weightage (min. value is 0.4 loading)
↳ sometimes 0.33

(i) Exploratory factor Analysis
(ii) Confirmatory " "

## * Rotations in Factor Analysis

↳ rotate the axes
↳ orthogonal rotation (90° bet. axes)
↳ oblique " (angle bet. ~~ang~~ axes less than 90°)



Disadvantage → Angle of rotation is subjective and this will change the interpretation.

# ITEM RESPONSE THEORY

## (Latent Response Theory)

CTT → Classical Test Theory
(achievement Test)

- Item difficulty
- Item discrimination (discriminate bet. people with good ability and those without it)
- Guessing (Pseudo)

* Binary Response → 1 (True)
→ 0 (False)

$x_1, x_2 \ldots x_f$ → items

$$\boxed{P(x_{ij} = 1) = \frac{e^{\eta_i - \alpha_j}}{1 + e^{\eta_i - \alpha_j}}}$$

**Rasch Model**

* $\eta_i$ → ability (latent trait) of subject $i$

$\alpha_j$ → difficulty parameter of item $j$

$x_{ij}$ → response of $i^{th}$ subject on $j^{th}$ item

when $\eta_i = \alpha_j$ ; Ability $= 0$

* Ogive Curve
→ Ability Factor
→ Discrimination Factor

* Item Characteristic Curve